sqoop import-all-tables --connect jdbc:mysql://quickstart.cloudera:3306/retail\_db --username retail\_dba --password cloudera --warehouse-dir /user/hive/warehouse/retail\_stage2.db --as-avrodatafile --compress --compression-codec org.apache.hadoop.io.compress.SnappyCodec -m 1

show databases;

create database retail\_stage2;

use retail\_stage2;

create table orders\_sqoop(order\_id int, order\_date bigint,order\_customer\_id int,order\_status string)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

STORED AS AVRO;

select os.order\_date,os.order\_id from orders\_sqoop os where os.order\_date IN (select T1.order\_date from(select order\_date,COUNT(distinct(order\_id)) num from orders\_sqoop group by order\_date order by num desc limit 1)T1);

select from\_unixtime(CAST(os.order\_date/1000 AS BIGINT)),os.order\_id from orders\_sqoop os where os.order\_date IN (select T1.order\_date from(select order\_date,COUNT(distinct(order\_id)) num from orders\_sqoop group by order\_date order by num desc limit 1)T1);

select from\_unixtime(CAST(os.order\_date/1000 AS BIGINT)),os.order\_id from orders\_sqoop os where os.order\_date IN (select T1.order\_date from(select oso.order\_date,COUNT(distinct(oso.order\_id)) num from orders\_sqoop oso group by oso.order\_date order by num desc limit 1)T1);

create table orders\_avro1(order\_id int,order\_date bigint,order\_customer\_id int, order\_status string)

partitioned by(order\_month string)

row format delimited fields terminated by ','

stored as avro;

insert into table orders\_avro1 PARTITION(order\_month) select order\_id,order\_date,order\_customer\_id,order\_status,substr(from\_unixtime(CAST((order\_date/1000) AS BIGINT)),1,7) as order\_month from retail\_stage2.orders\_sqoop;

select \* from orders\_avro1 oa where oa.order\_date IN(select T1.order\_date from (select oa1.order\_date,COUNT(distinct(oa1.order\_id)) num from orders\_avro1 oa1 group by oa1.order\_date order by num desc limit 1)T1);

ALTER TABLE orders\_avro1 ADD COLUMNS(order\_style string,order\_zone int);

insert into table orders\_sqoop values (99998,1383462000000,10540,'COMPLETE','Apparel',1),(99999,1383462000000,10540,'COMPLETE','Apparel',2);

select \* from orders\_sqoop os where os.order\_date IN(select T1.order\_date from(select os1.order\_date,COUNT(DISTINCT(os1.order\_id)) num from orders\_sqoop os1 group by os1.order\_date order by num desc limit 1)T1);

select to\_unix\_timestamp(from\_unixtime(CAST((order\_date/1000) AS BIGINT))) ord\_date from orders\_sqoop where order\_id =1;